

#6



OIPE

RAW SEQUENCE LISTING

DATE: 06/11/2002

PATENT APPLICATION: US/10/068,059

TIME: 14:45:57

Input Set : A:\SEQUENCE LISTING.TXT

Output Set: N:\CRF3\06112002\J068059.raw

ENTERED

C--> 4 <110> APPLICANT: Mizzen, Lee A.
5 Hongwei, Liu
6 Siegel, Marvin
8 <120> TITLE OF INVENTION: HEPATITIS B VIRUS TREATMENT
10 <130> FILE REFERENCE: 12071-017002
12 <140> CURRENT APPLICATION NUMBER: US 10/068,059
13 <141> CURRENT FILING DATE: 2002-06-04
15 <150> PRIOR APPLICATION NUMBER: US 60/266,733
16 <151> PRIOR FILING DATE: 2001-02-05
18 <160> NUMBER OF SEQ ID NOS: 12
20 <170> SOFTWARE: FastSEQ for Windows Version 4.0
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 558
24 <212> TYPE: DNA
25 <213> ORGANISM: Hepatitis B virus
27 <220> FEATURE:
28 <221> NAME/KEY: CDS
29 <222> LOCATION: (1)...(555)
31 <400> SEQUENCE: 1
32 atg gac att gac cct tat aaa gaa ttt gga gct act gtg gag tta ctc 48
33 Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu
34 1 5 10 15
36 tcg ttt ttg cct tct gac ttc ttt cct tcc gtc aga gat ctc cta gac 96
37 Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp
38 20 25 30
40 acc gcc tca gct ctg tat cgg gaa gcc tta gag tct cct gag cat tgc 144
41 Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys
42 35 40 45
44 tca cct cac cac acc gca ctc agg caa gcc att ctc tgc tgg ggg gaa 192
45 Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu
46 50 55 60
48 ttg atg act cta gct acc tgg gtg ggt aat aat ttg gaa gat cca gca 240
49 Leu Met Thr Leu Ala Thr Trp Val Gly Asn Asn Leu Glu Asp Pro Ala
50 65 70 75 80
52 tca agg gat cta gta gtc aat tat gtt aat act aac atg ggt tta aaa 288
53 Ser Arg Asp Leu Val Val Asn Tyr Val Asn Thr Asn Met Gly Leu Lys
54 85 90 95
56 att agg caa cta ttg tgg ttt cat ata tct tgc ctt act ttt gga aga 336
57 Ile Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg
58 100 105 110
60 gag act gta ctt gaa tat ttg gta tct ttc gga gtg tgg att cgc act 384
61 Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr
62 115 120 125

RAW SEQUENCE LISTING

DATE: 06/11/2002

PATENT APPLICATION: US/10/068,059

TIME: 14:45:57

Input Set : A:\SEQUENCE LISTING.TXT

Output Set: N:\CRF3\06112002\J068059.raw

```

64 cct cca gcc tat aga cca cca aat gcc cct atc tta tca aca ctt ccg      432
65 Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro
66      130                      135                      140
68 gaa act act gtt gtt aga cga cgg gac cga ggc agg tcc cct aga aga.      480
69 Glu Thr Thr Val Val Arg Arg Arg Asp Arg Gly Arg Ser Pro Arg Arg
70 145                      150                      155                      160
72 aga act ccc tcg cct cgc aga cgc aga tct caa tcg ccg cgt cgc aga      528
73 Arg Thr Pro Ser Pro Arg Arg Arg Arg Ser Gln Ser Pro Arg Arg Arg
74                      165                      170                      175
76 aga tct caa tct cgg gaa tct caa tgt tga      558
77 Arg Ser Gln Ser Arg Glu Ser Gln Cys
78      180                      185
80 <210> SEQ ID NO: 2
81 <211> LENGTH: 185
82 <212> TYPE: PRT
83 <213> ORGANISM: Hepatitis B virus
85 <400> SEQUENCE: 2
86 Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu
87 1      5      10      15
88 Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp
89      20      25      30
90 Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys
91      35      40      45
92 Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu
93      50      55      60
94 Leu Met Thr Leu Ala Thr Trp Val Gly Asn Asn Leu Glu Asp Pro Ala
95 65      70      75      80
96 Ser Arg Asp Leu Val Val Asn Tyr Val Asn Thr Asn Met Gly Leu Lys
97      85      90      95
98 Ile Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg
99      100     105     110
100 Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr
101     115     120     125
102 Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro
103     130     135     140
104 Glu Thr Thr Val Val Arg Arg Arg Asp Arg Gly Arg Ser Pro Arg Arg
105 145     150     155     160
106 Arg Thr Pro Ser Pro Arg Arg Arg Arg Ser Gln Ser Pro Arg Arg Arg
107     165     170     175
108 Arg Ser Gln Ser Arg Glu Ser Gln Cys
109     180     185
111 <210> SEQ ID NO: 3
112 <211> LENGTH: 513
113 <212> TYPE: DNA
114 <213> ORGANISM: Artificial Sequence
116 <220> FEATURE:
117 <221> NAME/KEY: CDS
118 <222> LOCATION: (1)...(510)
119 <223> OTHER INFORMATION: Nucleic acids encoding fusion protein

```

RAW SEQUENCE LISTING

DATE: 06/11/2002

PATENT APPLICATION: US/10/068,059

TIME: 14:45:57

Input Set : A:\SEQUENCE LISTING.TXT

Output Set: N:\CRF3\06112002\J068059.raw

```

121 <400> SEQUENCE: 3
122 atg ggc agc agc cat cat cat cat cat cac agc agc ggc ctg gtg ccg      48
123 Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
124   1           5           10           15
126 cgc ggc agc cat atg gac att gac cct tat aaa gaa ttt gga gct act      96
127 Arg Gly Ser His Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr
128           20           25           30
130 gtg gag tta ctc tcg ttt ttg cct tct gac ttc ttt cct tcc gtc aga      144
131 Val Glu Leu Leu Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg
132           35           40           45
134 gat ctc cta gac acc gcc tca gct ctg tat cgg gaa gcc tta gag tct      192
135 Asp Leu Leu Asp Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser
136           50           55           60
138 cct gag cat tgc tca cct cac cac acc gca ctc agg caa gcc att ctc      240
139 Pro Glu His Cys Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu
140   65           70           75           80
142 tgc tgg ggg gaa ttg atg act cta gct acc tgg gtg ggt aat aat ttg      288
143 Cys Trp Gly Glu Leu Met Thr Leu Ala Thr Trp Val Gly Asn Asn Leu
144           85           90           95
146 gaa gat cca gca tca agg gat cta gta gtc agt tat gtt aat act aac      336
147 Glu Asp Pro Ala Ser Arg Asp Leu Val Val Ser Tyr Val Asn Thr Asn
148           100          105          110
150 atg ggt tta aaa ttt agg caa cta ttg tgg ttt cat ata tct tgc ctt      384
151 Met Gly Leu Lys Phe Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu
152           115          120          125
154 act ttt gga aga gag act gta ctt gaa tat ttg gta tct ttc gga gtg      432
155 Thr Phe Gly Arg Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val
156           130          135          140
158 tgg att cgc act cct cca gcc tat aga cca cca aat gcc cct atc tta      480
159 Trp Ile Arg Thr Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu
160   145          150          155          160
162 tca aca ctt ccg gaa act act gtt gtt aac tga      513
163 Ser Thr Leu Pro Glu Thr Thr Val Val Asn
164           165          170
166 <210> SEQ ID NO: 4
167 <211> LENGTH: 170
168 <212> TYPE: PRT
169 <213> ORGANISM: Artificial Sequence
171 <220> FEATURE:
172 <223> OTHER INFORMATION: Fusion protein
174 <400> SEQUENCE: 4
175 Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
176   1           5           10           15
177 Arg Gly Ser His Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr
178           20           25           30
179 Val Glu Leu Leu Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg
180           35           40           45
181 Asp Leu Leu Asp Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser
182           50           55           60

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/068,059

DATE: 06/11/2002

TIME: 14:45:57

Input Set : A:\SEQUENCE LISTING.TXT

Output Set: N:\CRF3\06112002\J068059.raw

```

183 Pro Glu His Cys Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu
184 65 70 75 80
185 Cys Trp Gly Glu Leu Met Thr Leu Ala Thr Trp Val Gly Asn Asn Leu
186 85 90 95
187 Glu Asp Pro Ala Ser Arg Asp Leu Val Val Ser Tyr Val Asn Thr Asn
188 100 105 110
189 Met Gly Leu Lys Phe Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu
190 115 120 125
191 Thr Phe Gly Arg Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val
192 130 135 140
193 Trp Ile Arg Thr Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu
194 145 150 155 160
195 Ser Thr Leu Pro Glu Thr Thr Val Val Asn
196 165 170
198 <210> SEQ ID NO: 5
199 <211> LENGTH: 2241
200 <212> TYPE: DNA
201 <213> ORGANISM: Artificial Sequence
203 <220> FEATURE:
204 <221> NAME/KEY: CDS
205 <222> LOCATION: (1)...(2238)
206 <223> OTHER INFORMATION: Nucleic acids encoding fusion protein
208 <400> SEQUENCE: 5
209 atg ggc agc agc cat cat cat cat cat cac agc agc ggc ctg gtg ccg 48
210 Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
211 1 5 10 15
213 cgc ggc agc cat atg gac att gac cct tat aaa gaa ttt gga gct act 96
214 Arg Gly Ser His Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr
215 20 25 30
217 gtg gag tta ctc tcg ttt ttg cct tct gac ttc ttt cct tcc gtc aga 144
218 Val Glu Leu Leu Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg
219 35 40 45
221 gat ctc cta gac acc gcc tca gct ctg tat cgg gaa gcc tta gag tct 192
222 Asp Leu Leu Asp Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser
223 50 55 60
225 cct gag cat tgc tca cct cac cac acc gca ctc agg caa gcc att ctc 240
226 Pro Glu His Cys Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu
227 65 70 75 80
229 tgc tgg ggg gaa ttg atg act cta gct acc tgg gtg ggt aat aat ttg 288
230 Cys Trp Gly Glu Leu Met Thr Leu Ala Thr Trp Val Gly Asn Asn Leu
231 85 90 95
233 gaa gat cca gca tca agg gat cta gta gtc aat tat gtt aat act aac 336
234 Glu Asp Pro Ala Ser Arg Asp Leu Val Val Asn Tyr Val Asn Thr Asn
235 100 105 110
237 atg ggt tta aaa ttt agg caa cta ttg tgg ttt cat ata tct tgc ctt 384
238 Met Gly Leu Lys Phe Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu
239 115 120 125
241 act ttt gga aga gag act gta ctt gaa tat ttg gta tct ttc gga gtg 432
242 Thr Phe Gly Arg Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/068,059

DATE: 06/11/2002

TIME: 14:45:57

Input Set : A:\SEQUENCE LISTING.TXT

Output Set: N:\CRF3\06112002\J068059.raw

243	130	135	140	
245	tg	att	cgc	act cct cca gcc tat aga cca cca aat gcc cct atc tta 480
246	Trp	Ile	Arg	Thr Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu
247	145	150	155	160
249	tca	aca	ctt	cgc gaa act act gtt gtt aga cga cgg gac cga ggc agg 528
250	Ser	Thr	Leu	Pro Glu Thr Thr Val Val Arg Arg Arg Asp Arg Gly Arg
251		165	170	175
253	tcc	cct	aga	aga aga act ccc tcg cct cgc aga cgc aga tct caa tcg 576
254	Ser	Pro	Arg	Arg Arg Thr Pro Ser Pro Arg Arg Arg Arg Ser Gln Ser
255		180	185	190
257	ccg	cgt	cgc	aga aga tct caa tct cgg gaa tct caa tgt gtt aac gcc 624
258	Pro	Arg	Arg	Arg Arg Ser Gln Ser Arg Glu Ser Gln Cys Val Asn Ala
259		195	200	205
261	aag	aca	att	gcg tac gac gaa gag gcc cgt cgc ggc ctc gag cgg ggc 672
262	Lys	Thr	Ile	Ala Tyr Asp Glu Glu Ala Arg Arg Gly Leu Glu Arg Gly
263		210	215	220
265	ttg	aac	gcc	ctc gcc gat gcg gta aag gtg aca ttg ggc ccc aag ggc 720
266	Leu	Asn	Ala	Leu Ala Asp Ala Val Lys Val Thr Leu Gly Pro Lys Gly
267	225	230	235	240
269	cgc	aac	gtc	gtc ctg gaa aag aag tgg ggt gcc ccc acg atc acc aac 768
270	Arg	Asn	Val	Val Leu Glu Lys Lys Trp Gly Ala Pro Thr Ile Thr Asn
271		245	250	255
273	gat	ggt	gtg	tcc atc gcc aag gag atc gag ctg gag gat ccg tac gag 816
274	Asp	Gly	Val	Ser Ile Ala Lys Glu Ile Glu Leu Glu Asp Pro Tyr Glu
275		260	265	270
277	aag	atc	ggc	gcc gag ctg gtc aaa gag gta gcc aag aag acc gat gac 864
278	Lys	Ile	Gly	Ala Glu Leu Val Lys Glu Val Ala Lys Lys Thr Asp Asp
279		275	280	285
281	gtc	gcc	ggt	gac ggc acc acg acg gcc acc gtg ctg gcc cag gcg ttg 912
282	Val	Ala	Gly	Asp Gly Thr Thr Thr Ala Thr Val Leu Ala Gln Ala Leu
283		290	295	300
285	gtt	cgc	gag	ggc ctg cgc aac gtc gcg gcc ggc gcc aac ccg ctc ggt 960
286	Val	Arg	Glu	Gly Leu Arg Asn Val Ala Ala Gly Ala Asn Pro Leu Gly
287	305	310	315	320
289	ctc	aaa	cgc	ggc atc gaa aag gcc gtg gag aag gtc acc gac acc ctg 1008
290	Leu	Lys	Arg	Gly Ile Glu Lys Ala Val Glu Lys Val Thr Asp Thr Leu
291		325	330	335
293	ctc	aag	ggc	gcc aag gag gtc gag acc aag gag cag att gcg gcc acc 1056
294	Leu	Lys	Gly	Ala Lys Glu Val Glu Thr Lys Glu Gln Ile Ala Ala Thr
295		340	345	350
297	gca	gcg	att	tcg gcg ggt gac cag tcc atc ggt gac ctg atc gcc gag 1104
298	Ala	Ala	Ile	Ser Ala Gly Asp Gln Ser Ile Gly Asp Leu Ile Ala Glu
299		355	360	365
301	gcg	atg	gac	aag gtg ggc aac gag ggc gtc atc acc gtc gag gag tcc 1152
302	Ala	Met	Asp	Lys Val Gly Asn Glu Gly Val Ile Thr Val Glu Glu Ser
303		370	375	380
305	aac	acc	ttt	ggg ctg cag ctc gag ctc acc gag ggt atg ccg ttc gac 1200
306	Asn	Thr	Phe	Gly Leu Gln Leu Glu Leu Thr Glu Gly Met Arg Phe Asp
307	385	390	395	400

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/068,059

DATE: 06/11/2002

TIME: 14:45:58

Input Set : A:\SEQUENCE LISTING.TXT

Output Set: N:\CRF3\06112002\J068059.raw

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date